

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-169287

(P2001-169287A)

(43) 公開日 平成13年6月22日 (2001.6.22)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード* (参考)
H 0 4 N 7/30		H 0 3 M 7/30	A 5 C 0 5 9
H 0 3 M 7/30		H 0 4 N 7/133	Z 5 J 0 6 4
H 0 4 N 7/32		7/137	Z

審査請求 未請求 請求項の数16 O L (全 14 頁)

(21) 出願番号 特願2000-230768(P2000-230768)

(22) 出願日 平成12年7月31日(2000.7.31)

(31) 優先権主張番号 特願平11-281716

(32) 優先日 平成11年10月1日(1999.10.1)

(33) 優先権主張国 日本 (J P)

(71) 出願人 000005821  
松下電器産業株式会社  
大阪府門真市大字門真1006番地

(72) 発明者 井上 由紀子  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

(72) 発明者 有村 耕治  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

(74) 代理人 100097179  
弁理士 平野 一幸

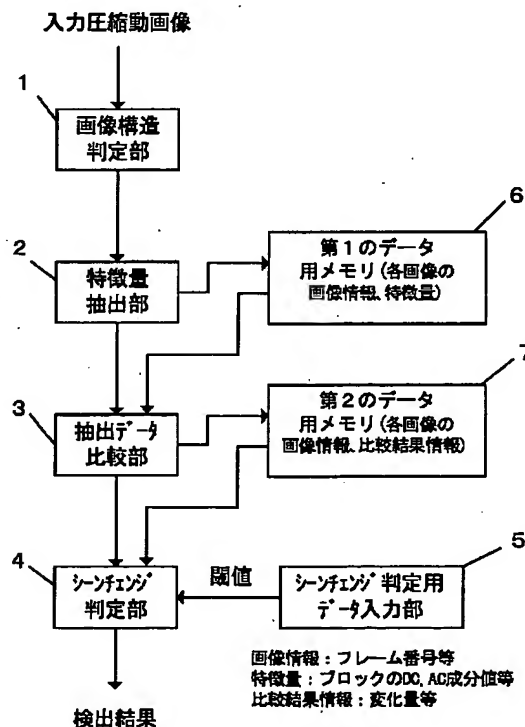
最終頁に続く

(54) 【発明の名称】 圧縮動画像のシーンチェンジ検出装置、圧縮動画像のシーンチェンジ検出方法及びそのプログラムを記録した記録媒体

(57) 【要約】

【課題】 フレーム/フィールド構造が混在してもシーンチェンジを検出可能な技術を提供する。

【解決手段】 入力圧縮動画像の画像構造を判定する画像構造判定部1と、画像構造判定部の判定結果がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に上下2倍分のデータに基づいて特徴量を抽出する特徴量抽出部2と、特徴量抽出部で抽出されたデータを記録するデータ用メモリ6と、抽出されたデータを比較して映像の変化量を求める抽出データ比較部3と、抽出データ比較部において求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定部4を有する。



## 【特許請求の範囲】

【請求項 1】入力圧縮動画像の画像構造を判定する画像構造判定部と、前記画像構造判定部の判定結果がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に上下 2 倍分のブロックデータに基づいて特徴量を抽出する特徴量抽出部と、前記特徴量抽出部で抽出されたブロックデータを記録する記憶領域と、抽出されたブロックデータを比較して映像の変化量を求める抽出データ比較部と、前記抽出データ比較部において求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定部を有することを特徴とする圧縮動画像のシーンチェンジ検出装置。

【請求項 2】入力圧縮動画像の構造画像に関係無く、1 ブロック分のブロックデータに基づいて特徴量を抽出する特徴量抽出部と、前記特徴量抽出部で抽出されたブロックデータを記録する記憶領域と、特徴量が抽出された画像がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に 2 倍のブロックデータを用いて特徴量を比較し変化量を求める抽出データ比較部と、前記抽出データ比較部において求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定部を有することを特徴とする圧縮動画像のシーンチェンジ検出装置。

【請求項 3】画像がフレーム構造画像である場合、フィールド DCT 符号化されたブロック数をカウントするフィールド DCT 符号化ブロック数カウント部と、このブロック数と閾値とを比較して、フィールド間に存在するシーンチェンジを判定するシーンチェンジ判定部とを有することを特徴とする圧縮動画像のシーンチェンジ検出装置。

【請求項 4】シーンチェンジを判定するシーンチェンジ判定部と、前記シーンチェンジ判定部が検出したシーンチェンジのうち、指定された特定の区間の開始点と終了点に存在するシーンチェンジを検索するシーンチェンジ区間検索部とを有することを特徴とする圧縮動画像のシーンチェンジ検出装置。

【請求項 5】前記シーンチェンジ判定部が、シーンチェンジの判定基準として用いる閾値には、画像の最大変化量を基準として定められる閾値が含まれていることを特徴とする請求項 1 から 4 記載の圧縮動画像のシーンチェンジ検出装置。

【請求項 6】フィールド構造画像とフレーム構造画像とが混在する圧縮動画像を入力し、入力した圧縮動画像におけるシーンチェンジを検出することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 7】入力圧縮動画像の画像構造を判定する画像構造判定ステップと、前記画像構造判定ステップの判定結果がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に上下 2 倍分のブロックデータに基づいて特徴量を抽出する特徴量抽出ステップと、前記

特徴量抽出ステップで抽出されたブロックデータを記録する記憶領域と、抽出されたブロックデータを比較して映像の変化量を求める抽出データ比較ステップと、前記抽出データ比較ステップにおいて求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定ステップを有することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 8】入力圧縮動画像の構造画像に関係無く、1 ブロック分のブロックデータに基づいて特徴量を抽出する特徴量抽出ステップと、前記特徴量抽出ステップで抽出されたブロックデータを記録する記憶領域と、特徴量が抽出された画像がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に 2 倍のブロックデータを用いて特徴量を比較し変化量を求める抽出データ比較ステップと、前記抽出データ比較ステップにおいて求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定ステップを有することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 9】画像がフレーム構造画像である場合、フィールド DCT 符号化されたブロック数をカウントするフィールド DCT 符号化ブロック数カウントステップと、このブロック数と閾値とを比較して、フィールド間に存在するシーンチェンジを判定するシーンチェンジ判定ステップとを有することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 10】シーンチェンジを判定するシーンチェンジ判定ステップと、前記シーンチェンジ判定ステップにおいて検出したシーンチェンジのうち、指定された特定の区間の開始点と終了点に存在するシーンチェンジを検索するシーンチェンジ区間検索ステップとを有することを特徴とする圧縮動画像のシーンチェンジ検出方法。

【請求項 11】前記シーンチェンジ判定ステップにおいて、シーンチェンジの判定基準として用いる閾値には、画像の最大変化量を基準として定められる閾値が含まれていることを特徴とする請求項 7 から 10 記載の圧縮動画像のシーンチェンジ検出方法。

【請求項 12】入力圧縮動画像の画像構造を判定する画像構造判定ステップと、前記画像構造判定ステップの判定結果がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に上下 2 倍分のブロックデータに基づいて特徴量を抽出する特徴量抽出ステップと、前記特徴量抽出ステップで抽出されたブロックデータを記録する記憶領域と、抽出されたブロックデータを比較して映像の変化量を求める抽出データ比較ステップと、前記抽出データ比較ステップにおいて求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定ステップを有することを特徴とする圧縮動画像のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【請求項 13】入力圧縮動画像の構造画像に関係無く、

10

20

30

40

50

1ブロック分のブロックデータに基づいて特徴量を抽出する特徴量抽出ステップと、前記特徴量抽出ステップで抽出されたブロックデータを記録する記憶領域と、特徴量が抽出された画像がフレーム構造画像である場合、フィールド構造画像に対して画像縦方向に2倍のブロックデータを用いて特徴量を比較し変化量を求める抽出データ比較ステップと、前記抽出データ比較ステップにおいて求められた変化量を用いてシーンチェンジを判定するシーンチェンジ判定ステップを有することを特徴とする圧縮動画のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【請求項14】画像がフレーム構造画像である場合、フィールドDCT符号化されたブロック数をカウントするフィールドDCT符号化ブロック数カウントステップと、このブロック数と閾値とを比較して、フィールド間に存在するシーンチェンジを判定するシーンチェンジ判定ステップとを有することを特徴とする圧縮動画のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【請求項15】シーンチェンジを判定するシーンチェンジ判定ステップと、前記シーンチェンジ判定ステップにおいて検出したシーンチェンジのうち、指定された特定の区間の開始点と終了点に存在するシーンチェンジを検索するシーンチェンジ区間検索ステップとを有することを特徴とする圧縮動画のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【請求項16】前記シーンチェンジ判定ステップにおいて、シーンチェンジの判定基準として用いる閾値には、画像の最大変化量を基準として定められる閾値が含まれていることを特徴とする請求項12から15記載の圧縮動画のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、圧縮動画から、シーンの変わり目を検出するシーンチェンジ検出装置及びその関連技術に関するものである。

【0002】

【従来の技術】近年、デジタルビデオカメラの普及や、デジタル放送の出現などにより、MPEGやDVなど、圧縮動画を扱う機会が増加している。また、過去の大量のアナログ映像を、デジタルの圧縮動画として保存する場合もある。そして、このような圧縮動画を復号せずに、符号化されたまま、編集する技術が実用化されつつある。

【0003】このような編集においては、圧縮動画、すなわちビットストリームのなかから、シーンチェンジ（映像又は場面の変り目の位置）を、高速に自動検出する技術が是非とも必要となる。なぜなら、検出されたシーンの先頭の位置情報やシーンチェンジ技術によって

切り分けられたシーンの代表画像は、映像内容のインデックスとして有用であり、内容の検索や編集の重要な手助けとなるからである。

【0004】次に、圧縮動画のフォーマットとして広く用いられている、MPEG2の符号化について説明する。MPEG2符号化は、MPEG1、H.261と同様に動きベクトルとDCT（離散コサイン変換：Discrete Cosine Transform）を用いた方式である。データは、フレーム内で、輝度（Y）、色差（Cb、Cr）に分離され、16\*16画素のマクロブロック単位で、符号化される。

【0005】そして、各マクロブロックを符号化するにあたっては、参照画像から動き予測を行う動き補償予測が選択されるか、または、符号化を行うデータのみで符号化を行うイントラ符号化が選択される。

【0006】このうち、動き補償予測は、フレーム間の時間的な相関が高い場合に符号化率が高くなる方式であり、符号化を行うマクロブロックと、参照画像から動き予測によって得られるマクロブロックのデータとの、差分から予測誤差信号を得て時間空間的に情報の圧縮を行うものである。なお、動き補償予測では、この予測誤差信号が、8\*8画素のブロック単位で、DCTにより空間周波数領域に変換される。

【0007】一方、イントラ符号化は、符号化されるブロックのデータそのものを、8\*8画素のブロックに分け、これらのブロック単位で、単純にDCT符号化する方式である。

【0008】以下、符号化の単位について説明する。MPEG2では、インタレース画像も対象としており、画面の符号化の単位として、フレーム構造とフィールド構造とがある。

【0009】フレーム構造では、奇数フィールドと偶数フィールドの2フィールドをインタレースに配置したフレームに対して符号化を行う。一方、フィールド構造では、奇数フィールドもしくは偶数フィールドの、1フィールドに対して符号化を行う。

【0010】さて、本明細書において、フレーム構造で符号化された画像を「フレーム構造画像」といい、フィールド構造で符号化された画像を「フィールド構造画像」という。

【0011】次に、動き補償について説明する。MPEG2では、上述のように、フレーム構造とフィールド構造とがある。そして、フレーム構造画像の動き補償予測については、フレーム予測、フィールド予測及びデュアルプライム（Dual-Prime）予測がある。また、フィールド構造画像の動き補償予測としては、フィールド予測、16\*8MC予測及びデュアルプライム予測がある。さらに、フレーム予測以外の予測においては、参照するフィールドが奇数フィールドであるか偶数フィールドであるかを選択することができる。

【0012】次に、図15を参照しながら、符号化の方法について説明する。フレーム構造画像では、符号化の方法として、フレームDCTとフィールドDCTという2種類のDCTを用いることができる。

【0013】このうち、フレームDCTは、マクロブロックの輝度信号を、図15(a)に示すように、4個のブロックに分解する際に、各ブロックがフレームで構成されるように分解し、これにDCTを施すものである。

【0014】一方、フィールドDCTは、図15(b)に示すように、マクロブロックの輝度信号を4個のブロックに分解する際に、各ブロックがフィールドで構成されるように分解し、これにDCTを施すものである。

【0015】さて、符号化の際には、この2種類のDCTのどちらを用いてもよいが、一般的に、奇数フィールドと偶数フィールドの画像データの差が大きい場合にフィールドDCTを用いると符号化の効率がよくなることが知られている。特に、二つのシーンが、一つのフィールドに混在しているような場合は、フィールドDCTを行うと圧縮率が高くなる。

【0016】しかし、フィールドDCTの場合、フレーム構造を2つのフィールドに分解する必要があるため、フレームDCTに対して処理速度は遅くなる。つまり、このような性質に合わせて、2種類のDCTをうまく用いることにより、フレーム構造画像（インタレース画像）の符号化効率を向上させることができる。なお、4:2:0フォーマットにおける色差信号については、常にフレームDCTを用いる。また、フィールド構造画像では、マクロブロックが1フィールドの信号のみで構成されるため、常にフィールドDCTを行うことになる。

【0017】以上の説明をふまえて、以下従来のシーンチェンジ技術について説明する。従来、シーンチェンジ検出には、(1)画像の色のヒストグラム、(2)圧縮動画のデータサイズ、(3)2フレームの画像の同一位置におけるブロックデータ差分、などの特徴量を用いている。

(1)画像の色のヒストグラムを用いるには、1フレームの画像に用いられている色を、1フレーム分あるいは1フレームを分割した領域において、ヒストグラムにし、このヒストグラムをそのフレームの特徴量とし、その前後のフレーム画像における特徴量と比較して類似度を求める（例えば、特開平7-59108号公報参照）。

(2)圧縮動画のデータサイズを用いるには、シーンチェンジ部分では、圧縮率が悪いという性質を利用して、隣り合うフレームのデータのサイズを比較し、その差分が所定の閾値よりも大きくなった時にシーンチェンジと判定する（例えば、特開平7-121555号公報参照）。

【0018】しかし、これら(1)～(2)の手法で

は、フレーム単位でしかシーンチェンジを検出できないから、シーンチェンジが、1フレーム中の奇数フィールドと偶数フィールドの間（2つのフィールド間）に存在すると、このシーンチェンジを正確に検出できない。

【0019】この問題を解決するために、特開平9-322120号公報では、フィールド予測方式を用いて符号化された映像データから復号処理を行うこと無しにシーンチェンジを検出する方法を提案している。このものでは、予測が行われるフレームにおいて、参照フレームの奇数フィールドもしくは偶数フィールドのどちらを選択して予測を行うかという、参照フィールド選択信号に基づいて、フィールド間の類似度を複数算出し、その結果を用いてシーンチェンジ検出を行っている。

【0020】

【発明が解決しようとする課題】しかしながら、この手法は、フィールド予測方式に依存しているから、フィールド予測方式（フレーム間予測方式）を用いない映像や、フィールド予測方式の映像とそれ以外の予測方式の映像が混在する映像に対しては、適用できない。

【0021】また、(3)データとして同じ位置のDCTにおけるDC成分のみの差分を用いると、位置の対応がとれない場合がある。これは、フレーム構造画像では、符号化の方法としてフレームDCTとフィールドDCTという2種類のDCTを用いることができるため、DCTを解かずにブロックデータの比較を行うと、比較する片方のデータがフレームDCTで符号化され、もう一方のデータがフィールドDCTで符号化されていた場合に、画像における8\*8画素のデータと、8\*16画素データ中の、奇数もしくは偶数フィールドのみの8\*8画素のデータとを、比較することになるからである。

【0022】これを改善するには、フレーム構造画像1フレーム分と、フィールド構造画像1フレーム相当分（奇数フィールド画像と偶数フィールド画像）とを比較しなければならない。しかし、これでは、2枚のフィールド構造画像のデータが揃った時点でしか比較ができず、処理が煩雑で処理速度も遅くなるという、問題点がある。

【0023】そこで本発明は、フィールド/フレームの各構造が混在していても、シーンチェンジを検出可能な技術を提供することを、第1の目的とする。

【0024】また、本発明は、フレーム構造のフィールド間に存在するシーンチェンジをも検出できる技術を提供することを、第2の目的とする。

【0025】さらに、本発明は、予め開始点と終了点との間隔が分かっている場合に、迅速に目的のシーンチェンジを検出できる技術を提供することを、第3の目的とする。

【0026】

【課題を解決するための手段】まず本発明では、フィールド構造画像とフレーム構造画像とが混在する圧縮動画

像を入力し、入力した圧縮動画像におけるシーンチェンジを検出する。そして本発明は、第1の目的のため、請求項1、2、6、8、12、13において、圧縮動画像がフレーム構造画像の場合には、フィールド構造画像に対応する縦方向に2倍のデータを抽出する。

【0027】本発明は、第2の目的のため、請求項3、9、14において、フィールドDCT符号化ブロック数カウンタ部を備える。

【0028】本発明は、第3の目的のため、請求項4、10、15において、複数の検出されたシーンチェンジの中から、指定された間隔に一致するシーンチェンジの組を検出し、そのシーンチェンジを結果として出力する。

【0029】

【発明の実施の形態】本発明におけるシーンチェンジ検出は、全て、予測方式に依存しない。したがって、予測方式が何であれ、あるいは、予測方式の有無にかかわらず、シーンチェンジを検出できる。なお以下、本発明の全ての実施の形態において、入力圧縮動画像は、MPEG2によることとして説明を行うが、DCTを用い、かつ、フィールド/フレームの各構造を混在させ得る圧縮方式であれば、同様の効果が得られるから、このような圧縮方式による技術であれば、本発明の対象に包含される。そして、本発明では、フィールド構造画像とフレーム構造画像とが混在する圧縮動画像を入力し、入力した圧縮動画像におけるシーンチェンジを検出するものである。

【0030】(第1の実施の形態)以下、図面を参照しながら、本発明の実施の形態を説明する。図1は、本発明の第1の実施の形態におけるシーンチェンジ検出装置のブロック図である。

【0031】図1に示すように、このシーンチェンジ検出装置には、MPEG2に従って符号化されたビットストリームが、入力圧縮動画像として入力される。また、このシーンチェンジ検出装置からは、ビットストリームにおいて検出されたシーンチェンジの位置を示す情報が、検出結果として出力される。ここで、従来の技術の項で述べたように、入力圧縮動画像の符号化方式、符号化の単位(フレーム/フィールド構造)、DCTなどは、種々のものが考えられ、また、1種類だけでなく、複数種のものが、時間軸上で混在していてもよい。

【0032】就中、フィールド構造画像とフレーム構造画像とが、時間軸上で交互に存在するような場合でも差し支えない。以上の点は、後述する他の実施の形態についても、そのまま適合する。

【0033】図1に示すように、入力圧縮動画像は、まず、画像構造判定部1に入力される。画像構造判定部1は、現在入力している画像が、フィールド構造画像、フレーム構造画像のいずれであるかを、ビットストリームの特定領域の情報を参照して、判定する。そして、この

判定結果と、ビットストリームの内容が、次段の特徴量抽出部2に出力される。

【0034】次に、特徴量、変化量及び特徴量抽出部2の動作を、図5～図6を例にとって、説明する。図5では、図の左側が時間軸上で古い画像を示し、右側が新しい画像を示し、 $t$ 枚目の画像を、画像 $t$ のように記述する。また、この例では、画像 $t-2 \sim t+1$ 、 $t+4 \sim t+5$ がフィールド構造画像であり、画像 $t+2 \sim t+3$ はフレーム構造画像である。即ち、画像 $t+1 \sim t+2$ 、 $t+3 \sim t+4$ では、構造の変化がある。

【0035】そして、図6(a)(トップフィールド)、図6(b)(ボトムフィールド)のような、フィールド構造画像では、1ブロック分のブロックデータ $D_t$ を用い、ブロックデータ $D_t$ をそのまま特徴量 $d_t$ として、使用する( $d_t = D_t$ )。

【0036】また、図6(c)のような、フレーム構造画像では、上下2ブロック分のブロックデータ(ブロックデータ $D_{tu}$ とブロックデータ $D_{tb}$ )を用い、その平均値を特徴量 $d_t$ として使用する( $d_t = (D_{tu} + D_{tb}) / 2$ )。

【0037】ここで、ブロックデータ $D_t$ 、 $D_{tu}$ 、 $D_{tb}$ は、同種のデータでなければならないが、ブロック内の輝度平均値やその他ブロック内の画像を代表する種々のデータを用いることができる。

【0038】また、図6(a)、(b)、(c)に示しているように、フレーム構造画像は、フィールド構造画像に対して、上下2倍の高さを持つので、フィールド構造画像とフレーム構造画像とを比較するには、フィールド構造画像において、ブロック座標 $(x, y)$ のブロックデータ $D_t$ と、ブロック座標 $(x, 2 \cdot y)$ のブロックデータ $D_{tu}$ 及びブロック座標 $(x, 2 \cdot y + 1)$ のブロックデータ $D_{tb}$ との組とを、対比して用いる。

【0039】次に、時間軸上で異なる位置にある、画像 $t$ と画像 $s$ (典型的には、 $s = t + 1$ )とにおける、ブロック座標 $(x, y)$ の特徴量を、それぞれ $d_t(x, y)$ 、 $d_s(x, y)$ とすると、画像 $t$ と画像 $s$ とにおける、第1の変化量 $R(t, s)$ を次のように定義する。この第1の変化量 $R(t, s)$ は、画像 $t$ と画像 $s$ との単純な差分に近い。

【0040】

【数1】

$$R(t, s) = \sum_{x, y} |d_t(x, y) - d_s(x, y)|$$

【0041】また、同様に、第2の変化量 $Q(t, s)$ を次のように定義する。第2の変化量 $Q(t, s)$ は、画像 $t$ と画像 $s$ の微分値に近い。

【0042】

【数2】

$$Q(t, s) = 2R(t, s) - R(t-1, t) - R(s, s+1)$$

【0043】ここで、第1の変化量 $R(t, s)$ と、これに対して、予め経験的に設定される第1の閾値とを、比較することにより、画像に大きな変化が現れた位置を取り出すことができ、この位置をシーンチェンジと検出できる。但し、画像によっては、動きの激しいものなど、シーンは同じではあるが、連続的な変化が起き、第1の変化量 $R(t, s)$ のみでは、過剰にシーンチェンジを検出するおそれがある。このような場合、第2の変化量 $Q(t, s)$ と、これに対して、予め設定される第2の閾値との、比較を併用することにより、不連続で、かつ、大きな変化がある位置のみを、シーンチェンジと検出することができ、検出されるシーンチェンジの信頼性を向上できる。

【0044】なお、動画像によっては、第1の変化量 $R(t, s)$ 、第2の変化量 $Q(t, s)$ の一方のみで十分な場合もあり得るし、第1の変化量 $R(t, s)$ 、第2の変化量 $Q(t, s)$ に適当な重みづけをして、両方検討するようにしても差し支えない。勿論、以上の2つの式は、単なる例示に過ぎず、これらと等価な式はもとより、シーンチェンジを判定できるものであれば、他の式で代用しても良い。

【0045】これらの第1の変化量 $R(t, s)$ 、第2の変化量 $Q(t, s)$ を求めるには、図3に示すように、入力画像1枚分の抽出データの全てを得た後に、比較対象画像との比較をし変化量を算出しても良いが、図4に示すように、1ブロック（もしくは複数ブロック）のデータを抽出した時に、同時に、比較対象画像の同じ位置に対応するデータとの比較を行う方が望ましい。なぜなら、図4のようにした方が、処理速度は上がるからである。

【0046】このようにすると、時間軸上において、フィールド構造画像とフレーム構造画像の混合する一連の映像から、一律にシーンチェンジを検出できる。

【0047】さて、図1に示すように、特徴量抽出部2が抽出した特徴量 $d_t$ と、特徴量 $d_t$ が抽出されたフレーム番号などの画像情報とが、第1のデータメモリ6に対応付けて格納される。この第1のデータメモリ6は、記憶領域に相当するものであり、ハードディスク装置などの他の記憶手段で代用しても良い。

【0048】特徴量抽出部2の次段にある抽出データ比較部3は、特徴量抽出部2が新たに特徴量 $d_t$ を抽出し、第1のデータメモリ6に、この新たな特徴量 $d_t$ を記録すると、この新たな特徴量 $d_t$ と、それ以前（典型的にはその直前回）に第1のデータメモリ6に記録された特徴量とを参照し、これらの特徴量間の変化量を求める。この変化量は、第1の変化量 $R(t, s)$ 、第2の変化量 $Q(t, s)$ の一方のみでも良いが、望ましくは、両方求める。

【0049】そして、抽出データ比較部3は、求めた変化量を比較結果情報として、第2のデータメモリ7に

記録すると共に、この変化量が生じたフレーム番号等の画像情報を、比較結果情報に対応付けて第2のデータメモリ7に記録する。

【0050】また、シーンチェンジ判定用データ入力部5は、シーンチェンジ判定時に使用する閾値を保持し、この閾値をシーンチェンジ判定部4へ出力する。この閾値としては、第1の変化量 $R(t, s)$ 用の第1の閾値と、第2の変化量 $Q(t, s)$ 用の第2の閾値とがある。

10 【0051】シーンチェンジ判定部4は、抽出データ比較部3が比較を行うと、第2のデータメモリ7に記録された比較結果情報を参照し、これがシーンチェンジ判定用データ入力部5から入力する閾値を越えていれば、この位置でシーンチェンジが発生したと判定し、この比較結果情報に係る、フレーム番号等の画像情報（つまり、ビットストリーム内の位置）を検出結果として出力する。また、そうでなければ、シーンチェンジ判定部4は、検出結果を出力しないようにしても良いし、あるいは、検出しなかった旨の検出結果を出力するようにしても良い。

30 【0052】図2は、本発明の第1の実施の形態におけるシーンチェンジ検出装置のフローチャートである。次に、図2を用いて、本形態のシーンチェンジ検出装置の動作を説明する。まず、入力圧縮動画像が、画像構造判定部1に至ると、この判定部1は、現在の画像が、フレーム構造画像／フィールド構造画像のいずれであるかを判定する（ステップ1）。フレーム構造画像であれば、特徴量抽出部2は、画像の縦方向上下2ブロック分のデータを用いた値を特徴量として抽出し、第1のデータメモリ6に記録する（ステップ2）。フィールド構造画像であれば、特徴量抽出部2は、1ブロック分のデータを用いた値を特徴量として抽出し、第1のデータメモリ6に記録する（ステップ3）。つまり、フレーム構造画像ではフィールド構造画像の縦方向2倍のデータを用いる。

40 【0053】そして、ステップ4にて、抽出データ比較部3が、今回の特徴量と、それ以前の特徴量とを比較して、比較結果情報を第2のデータメモリ7に記録する。次に、ステップ5にて、シーンチェンジ判定部4は、この比較結果情報をシーンチェンジ判定用データ入力部5から入力する、閾値と比較し、シーンチェンジと判定できれば、シーンチェンジが発生した位置を検出結果として出力する。

50 【0054】なお、本形態では、特徴量 $d_t$ として、フィールド構造画像では、1ブロック分のブロックデータ $D_t$ を用い、フレーム構造画像では、縦方向に2ブロック分のブロックデータ $D_{tu}$ 、 $D_{tb}$ の平均値を用いたが、フィールド構造画像とフレーム構造画像のデータの比較レベルが同等になれば、他の手法によっても良い。例えば、フィールド構造画像の1ブロック分のブロック



データを2倍にしたもの( $d_t = 2 * D_t$ )と、フレーム構造画像の縦方向上下2ブロックのブロックデータの和( $d_t = D_{tu} + D_{tb}$ )とを、特徴量 $d_t$ として、これらを比較してもよい。

【0055】また、本形態では、フィールド構造画像において抽出されるデータを1ブロックずつの単位とし、フレーム構造画像において抽出されるデータを2ブロックずつの単位としたが、フレーム構造画像から抽出されるデータが、フィールド構造画像において抽出されるデータと対応する位置に存在する縦方向に2倍のデータであれば、同様にシーンチェンジを検出できる。つまり、フィールド構造画像Nブロックのデータの特徴量として抽出した場合、フレーム構造画像からは対応する位置の縦方向に2倍の、 $2 * N$ ブロックからのデータを、特徴量として抽出するようにすればよい。

【0056】これにより、フレーム構造画像とフィールド構造画像の混合する圧縮動画像からも、一律にシーンチェンジを検出できる。また、フレーム構造画像内でフレームDCTを用いられようとフィールドDCTを用いられようと、特別な配慮をすることなくシーンチェンジを検出できる。また、この手法は、予測方式に依存しないので、予測方式が何であるかは不問であるし、予測方式の有無にかかわらず、所望のシーンチェンジを検出できる。

【0057】(第2の実施の形態)図7は、本発明の第2の実施の形態におけるシーンチェンジ検出装置のブロック図、図8は、本発明の第2の実施の形態におけるシーンチェンジ検出装置のフローチャートである。

【0058】図7では、第1の実施の形態に係る図1に対し、特徴量抽出部2と画像構造判定部1の位置を入れ換えてある。また、本形態の特徴量抽出部2は、入力された圧縮動画像がフレーム構造画像であってもフィールド構造画像であっても、画面全体について、1ブロックずつのブロックデータから、特徴量を抽出する点異なる(ステップ10)。

【0059】そして、図8に示すように、抽出データ比較部3の前段にある、画像構造判定部1は、今回入力した画像の構造を調べ、フレーム構造画像の場合、抽出データ比較部3は、画像の縦方向に上下2ブロック分のデータを用いてデータを比較する(ステップ12)。一方、フィールド構造画像ならば、抽出データ比較部3は、1ブロック分のデータを用いてデータを比較する(ステップ13)。ここでの比較は、第1の実施の形態と同様である。

【0060】さて、図5の例でいえば、画像 $t$ と画像 $t+1$ とを比較するとき、両画像ともフィールド構造画像であるので、画面上の同じ位置にあるブロックのデータが比較される。

【0061】画像 $t+1$ と画像 $t+2$ の比較であれば、画像 $t+2$ はフレーム構造画像であるので、画像 $t+1$

におけるブロック座標( $x, y$ )から得られるデータと、画像 $t+2$ のデータは、ブロック座標( $x, 2 * y$ )とブロック座標( $x, 2 * y + 1$ )から得られるデータとが、比較される。

【0062】そして、例えば、画像 $t+1$ におけるブロック座標( $x, y$ )から得られるデータをA、ブロック座標( $x, 2 * y$ )とブロック座標( $x, 2 * y + 1$ )から得られるデータをそれぞれB、Cとすると、データAとデータB、Cの平均値との差分の絶対値を、変化量とする。この変化量を画像全体において求めることで、2枚の画像の変化量が得られる。

【0063】なお、ここでは、フレーム構造画像のデータとして縦方向に2ブロック分の平均値を用いたが、これは単なる和でも良く、その場合、対応するフィールド構造画像のデータは、2倍にするなど、データの正規化を行えば良い。具体的には、データAの2倍の値とデータB、Cの和との差分の絶対値を変化量とすると良い。

【0064】その他の点は、第1の実施の形態と同様である。

【0065】(第3の実施の形態)図9は、本発明の第3の実施の形態におけるシーンチェンジ検出装置のブロック図、図10は、本発明の第3の実施の形態におけるシーンチェンジ検出装置のフローチャートである。

【0066】本形態では、図1と図9とを比較すれば明らかのように、フィールドDCT符号化ブロック数カウント部8と、第3のデータ用メモリ9を追加している。このフィールドDCT符号化ブロック数カウント部8は、画像構造判定部1による判定結果が、フレーム構造画像であった場合、その符号化がフレームDCTを用いているのかフィールドDCTを用いているのかを判定し、1フィールド中のフィールドDCT符号化が行われているマクロブロック数(ブロック数でも良い)をカウントする。そして、フィールドDCT符号化ブロック数カウント部8がカウントしたフィールドDCT符号化ブロック数は、第3のデータ用メモリ9に記録される。

【0067】したがって、図10に示すように、まず、入力圧縮動画像が、画像構造判定部1に至ると、この判定部1は、現在の画像が、フレーム構造画像/フィールド構造画像のいずれであるかを判定する(ステップ20)。フレーム構造画像であれば、特徴量抽出部2は、画像の縦方向上下2ブロック分のブロックデータを用いた値を特徴量として抽出し、第1のデータ用メモリ6に記録し(ステップ21)、フィールドDCT符号化ブロック数カウント部8がフィールドDCT符号化ブロック数をカウントし、このブロック数が第3のデータ用メモリ9に記録される(ステップ23)。なお、ステップ21、23の順序は入れ換えても差し支えない。

【0068】一方、フィールド構造画像であれば、特徴量抽出部2は、1ブロック分のブロックデータを用いた値を特徴量として抽出し、第1のデータ用メモリ6に記

録する(ステップ22)。つまり、フレーム構造画像ではフィールド構造画像の縦方向2倍のデータを用いる。

【0069】そして、ステップ24にて、抽出データ比較部3が、今回の特徴量と、それ以前の特徴量とを比較して、比較結果情報を第2のデータ用メモリ7に記録する。次に、ステップ25にて、シーンチェンジ判定部4は、この比較結果情報をシーンチェンジ判定用データ入力部5から入力する、閾値と比較し、シーンチェンジと判定できるかどうか検討する。さらに、ステップ26にて、シーンチェンジ判定部4は、第3のデータ用メモリ9に記録した、ブロック数と閾値とを比較して、シーンチェンジと判定できるかどうか検討する。そして、シーンチェンジ判定部4は、ステップ25又はステップ26のいずれかで、シーンチェンジと判定したら、発生した位置を検出結果として出力する。

【0070】さて、図11のように、二つのシーンがフィールドで混ざっているような場合には、入力画像と直前直後の画像との変化量が小さくなってしまう場合があり、検出漏れを起こす原因となっていた。しかし、従来の技術の項で述べたように、このような入力画像においてフィールドDCTを行うと圧縮率が高くなる。

【0071】そのため、フレーム内でフィールドDCTが多く用いられている場合には、フレーム内の奇数フィールドと偶数フィールドの相関が低いと見做すことができる。このため、特徴量抽出部2は、フィールドDCT符号化が行われている数をカウントし、これをフレーム内の第3の変化量として、比較検討対象に追加する。

【0072】抽出データ比較部3は、第1、第2の実施の形態と同様であるが、シーンチェンジ判定部4は、第1、第2の実施の形態における判定に加えて、第3の変化量と、この第3の変化量のために予め設定された、第3の閾値とを比較した場合に、第3の変化量が第3の閾値よりも大きい時には、フレームのフィールド間にシーンチェンジがあると判断する。

【0073】以上、説明したように、フィールドDCTが用いられた数をカウントすることにより、従来検出が非常に困難であった、フレーム構造画像の2つのフィールド間に存在するシーンチェンジ(図11に例示している)を検出できる。

【0074】さらに、本形態では、シーンチェンジ判定用データ入力部5が、シーンチェンジ判定部4に出力する閾値について、次の工夫がなされている。即ち、画像の最大変化量を基準(100%)として、閾値は、この基準の所定パーセントと定める。

【0075】例えば、比較画像A、Bにおいて、比較する1ブロックのデータの取り得る最小値が0であり最大値が255であるとする、1ブロックの最大変化量は255である。そして、画像の比較に用いたブロック数が1320であるとする、画像全体の最大変化量は、 $255 \times 1320 = 336600$ となる。この変化量を

基準(100%)とする。また、ここでの閾値は、例えば3%(10098)~10%(33660)程度が好適である。

【0076】勿論、使用するデータの数やデータの最大変化量が変われば、それに伴って閾値は変化するが、閾値と基準の比率は一定とする。

【0077】これにより、画像サイズ(縦×横)が変わったり、判定に使用されるデータの種類が変わったりしても、検出のばらつきを抑制して、ほぼ様な検出結果を得ることができる。

【0078】(第4の実施の形態)図12は、本発明の第4の実施の形態におけるシーンチェンジ検出装置のブロック図、図13は、本発明の第4の実施の形態におけるシーンチェンジ検出装置のフローチャートである。

【0079】本形態では、図1と図12とを比較すれば明かなように、シーンチェンジ判定部4が検出結果をダイレクトに出力するのではなく、シーンチェンジ判定部4がシーンチェンジと判定したシーンチェンジ位置情報を、一旦、第4のデータ用メモリ11に格納するようにしている。

【0080】また、シーンチェンジ判定用データ入力部5には、基準となるシーンチェンジからターゲットシーンまでの時間軸上の間隔が設定される。そして、シーンチェンジ間隔検索部10を追加している。このシーンチェンジ間隔検索部10は、第4のデータ用メモリ11に記録されたシーンチェンジ位置情報同士の時間軸上の間隔を求め、求めた間隔と、シーンチェンジ判定用データ入力部5から与えられる間隔とを、比較する。

【0081】したがって、図13に示すように、先の実施の形態と同様に、シーンチェンジ判定部4は、シーンチェンジを探す(ステップ30)。そして、シーンチェンジ判定部4が、シーンチェンジを見つけると、このシーンチェンジ位置情報を第4のデータ用メモリ11に格納する(ステップ31)。

【0082】そして、シーンチェンジ間隔検索部10は、第4のデータ用メモリ11をアクセスして、シーンチェンジ間の間隔を調べ(ステップ32)、シーンチェンジ判定用データ入力部5から与えられた間隔と一致するシーンチェンジの組が見つかる、見つかったシーンチェンジの先頭と末尾とからなる、組の位置情報を、検出結果として出力する(ステップ33)。

【0083】例えば、5分の映像の中から、30秒のシーンを検出したい場合、映像全体からシーンチェンジ検出を行い、その結果を、第4のデータ用メモリ11に記録する。その後、記録されたデータの中から、ちょうど30秒間隔になっているシーンチェンジの組を探出し、それを検出結果として出力する。

【0084】図14を例にとると、シーンチェンジ1~シーンチェンジ4が検出された場合に、シーンチェンジ1とシーンチェンジ4がちょうど30秒間隔であれば、



シーンチェンジ1とシーンチェンジ4の組が、検出結果として出力される。

【0085】このことにより、例えば、テレビ放送などから得られた大量な映像からCM部分のみを取り出したり、放送時間長の決まったニュースや番組を取り出したることができる。

【0086】さらに、あるシーンチェンジが見つかった場合、それから与えられた間隔までのシーンチェンジ判定を省略でき、無駄な検出動作を極力省いて、処理時間を短縮できる。

【0087】例えば、長い映像の中から、15秒のCMだけを検出したい場合には、ターゲット時間として15秒を与える。シーンチェンジ間隔検索部10では、演算によってシーンチェンジとして検出されたフレームから、15秒後のフレームにシーンチェンジが検出されるかを判定し、15秒後にシーンチェンジが検出された場合にのみ、そのフレームと15秒後のフレームを要求されたシーンチェンジとして出力する。この15秒間がターゲットのシーンとして検出されることになる。続けてターゲットシーンを検索するには、最後に検出されたシーンチェンジからまた15秒後にシーンチェンジが存在するかを判定し、シーンチェンジが検出されなければ次のフレームからシーンチェンジ判定を継続する。検出されればそこがターゲットシーンとなる。

【0088】図14を例にとると、シーンチェンジ1が検出され、そのちょうど15秒後にシーンチェンジ4があったとすると、この15秒間に存在する、シーンチェンジ2とシーンチェンジ3との判定（無駄な判定）をスキップすることができる。

【0089】このように、ここでは、一つ目のシーンチェンジが見つかったから、ターゲットの間隔として与えられた時間だけ後のフレームがシーンチェンジであるかどうかを判定している。したがって、検出されたターゲットシーンの中にあるシーンチェンジ判定のための処理を省くことができ、処理時間を短縮できる。

【0090】ここで、本明細書にいう「圧縮動画像のシーンチェンジ検出プログラムをコンピュータ読み取り可能に記録した記録媒体」には、複数の記録媒体にプログラムを分散して配布する場合を含む。また、このプログラムが、オペレーティングシステムの一部であるか否かを問わず、種々のプロセスないしスレッド（DLL、OCX、ActiveX等（マイクロソフト社の商標を含む））に機能の一部を肩代わりさせている場合には、肩代わりさせた機能に係る部分が、記録媒体に格納されていない場合も含む。

【0091】図1、図7、図9及び図12（以下「図1等」という）には、スタンドアロン形式のシステムを例示したが、サーバー／クライアント形式にしても良い。つまり、1つの端末機のみ、本明細書に出現する全ての要素が含まれている場合の他、1つの端末機がクライ

アントであり、これが接続可能なサーバないしネットワーク上に、全部又は一部の要素が実存していても、差し支えない。

【0092】さらには、図1等のほとんどの要素をサーバー側で持ち、クライアント側では、例えば、WWWブラウザだけにしても良い。この場合、各種の情報は、通常サーバ上にあり、基本的にネットワークを経由してクライアントに配布されるものだが、必要な情報が、サーバ上にあるときは、そのサーバの記憶装置が、ここにいう「記録媒体」となり、クライアント上にあるときは、そのクライアントの記録装置が、ここにいう「記録媒体」となる。

【0093】さらに、この「圧縮動画像のシーンチェンジ検出プログラム」には、コンパイルされて機械語になったアプリケーションの他、上述のプロセスないしスレッドにより解釈される中間コードとして実存する場合や、少なくともリソースとソースコードとが「記録媒体」上に格納され、これらから機械語のアプリケーションを生成できるコンパイラ及びリンカが「記録媒体」にある場合や、少なくともリソースとソースコードとが「記録媒体」上に格納され、これらから中間コードのアプリケーションを生成できるインタープリタが「記録媒体」にある場合なども含む。

【0094】

【発明の効果】本発明では、フィールド構造画像とフレーム構造画像とが混在する圧縮動画像を入力し、入力した圧縮動画像におけるシーンチェンジを検出するものである。したがって、画像構造の如何を気にすることなく、シーンチェンジを検出でき、圧縮動画像を検索する際の、重要なインデックスを得ることができる。そして、請求項1、2、7、8、12、13の構成によれば、フレーム構造画像とフィールド構造画像が時間軸上で混在していても、一律にシーンチェンジを検出できる。また、フレーム構造画像内でフレームDCTを用いられようとフィールドDCTを用いられようと、特別な配慮なしに、シーンチェンジを検出できる。

【0095】請求項3、9、14の構成によれば、フィールド符号化ブロック数をカウントし、このカウント数と閾値とを比較することによって、従来検出が非常に困難であった、フィールド間に存在するシーンチェンジをも検出できる。

【0096】請求項4、10、15の構成によれば、長い入力動画像から一部のターゲットを容易に取り出し得るし、無駄なシーンチェンジ判定を極力省略でき、その結果、処理時間を短縮できる。

【0097】請求項5、11、16の構成によれば、画像の大きさや、データの種類が変わっても、検出のばらつきが少なく、一様の検出結果を得ることができる。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態におけるシーンチェ

10

20

30

40

50

ンジ検出装置のブロック図

【図2】本発明の第1の実施の形態におけるシーンチェンジ検出装置のフローチャート

【図3】同詳細フローチャート

【図4】同詳細フローチャート

【図5】本発明の第1の実施の形態における入力圧縮動画像のモデル図

【図6】(a) 本発明の第1の実施の形態におけるフィールド構造画像の例示図(トップフィールド) (b) 同フィールド構造画像の例示図(ボトムフィールド)

(c) 同フレーム構造画像の例示図

【図7】本発明の第2の実施の形態におけるシーンチェンジ検出装置のブロック図

【図8】本発明の第2の実施の形態におけるシーンチェンジ検出装置のフローチャート

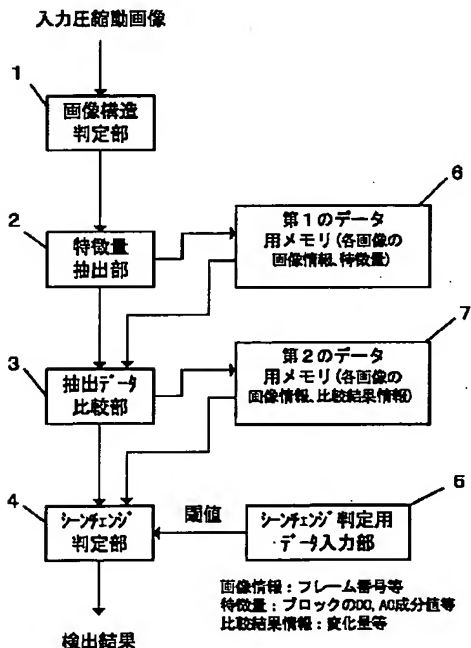
【図9】本発明の第3の実施の形態におけるシーンチェンジ検出装置のブロック図

【図10】本発明の第3の実施の形態におけるシーンチェンジ検出装置のフローチャート

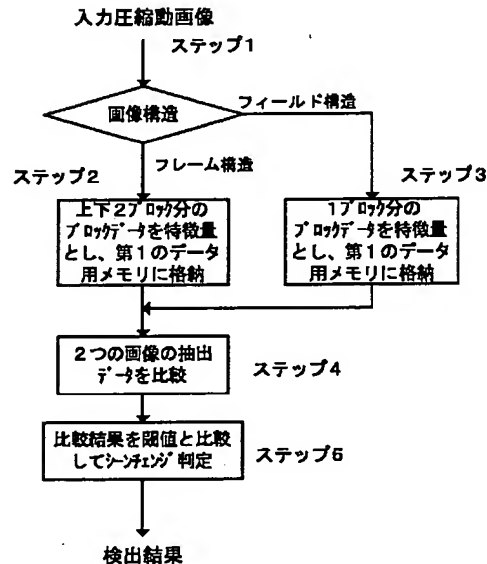
【図11】本発明の第3の実施の形態におけるフィールド間に存在するシーンチェンジの例示図

【図12】本発明の第4の実施の形態におけるシーンチ\*

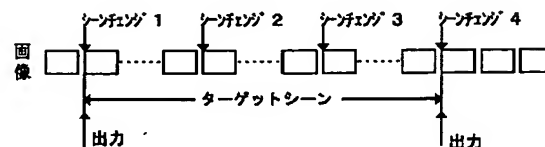
【図1】



【図2】



【図14】



\* エンジン検出装置のブロック図

【図13】本発明の第4の実施の形態におけるシーンチェンジ検出装置のフローチャート

【図14】本発明の第4の実施の形態におけるシーンチェンジのモデル図

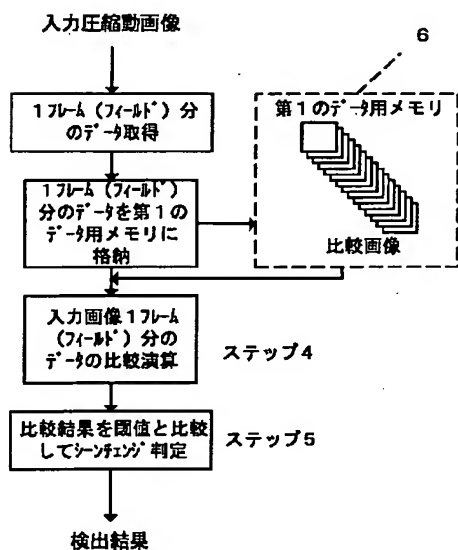
【図15】(a) 従来のフレームDCTにおけるマクロブロック構造の説明図

(b) 従来のフィールドDCTにおけるマクロブロック構造の説明図

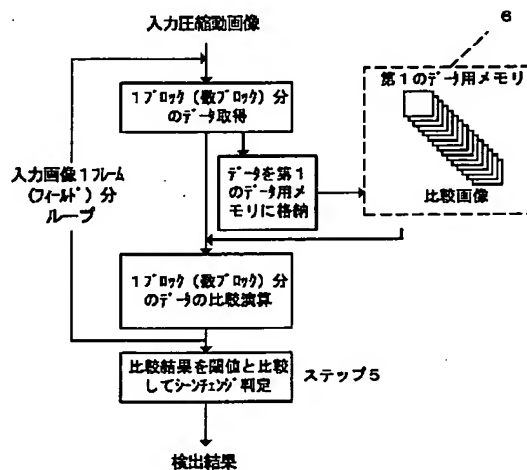
10 【符号の説明】

- 1 画像構造判定部
- 2 特徴量抽出部
- 3 抽出データ比較部
- 4 シーンチェンジ判定部
- 5 シーンチェンジ判定用データ入力部
- 6 第1のデータ用メモリ
- 7 第2のデータ用メモリ
- 8 フィールドDCT符号化ブロック数カウント部
- 9 第3のデータ用メモリ
- 10 シーンチェンジ間隔検索部
- 11 第4のデータ用メモリ

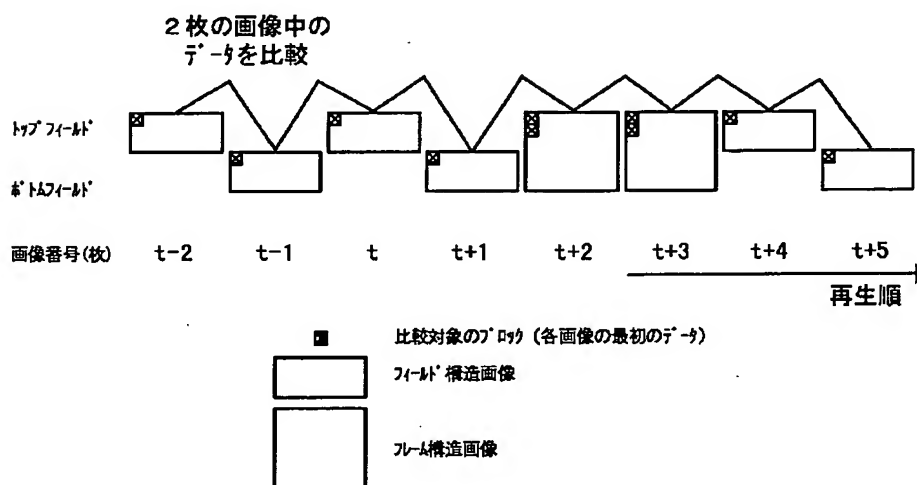
【図3】



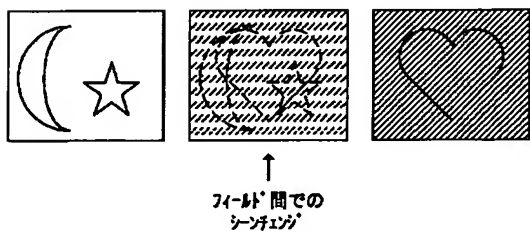
【図4】



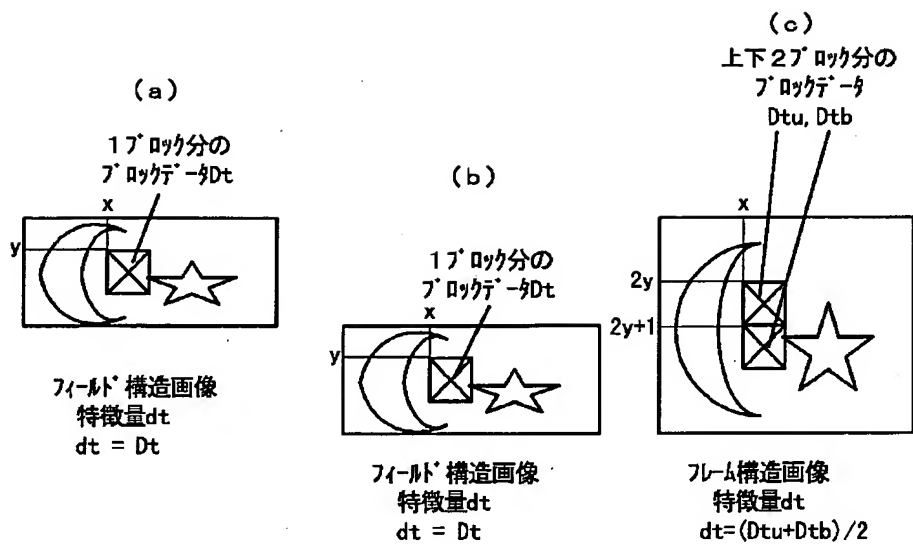
【図5】



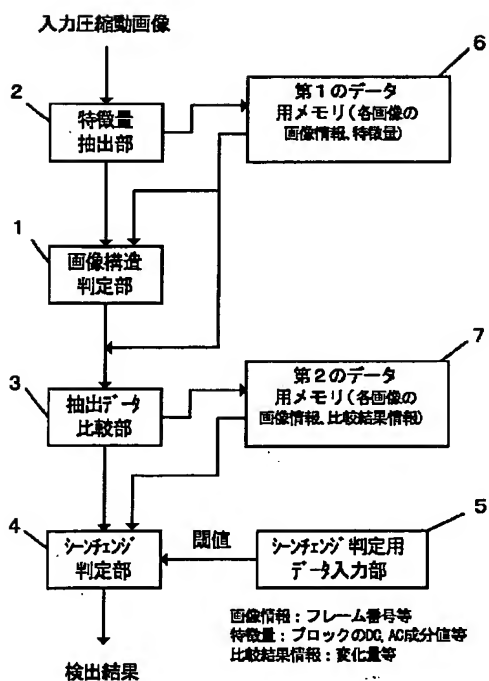
【図11】



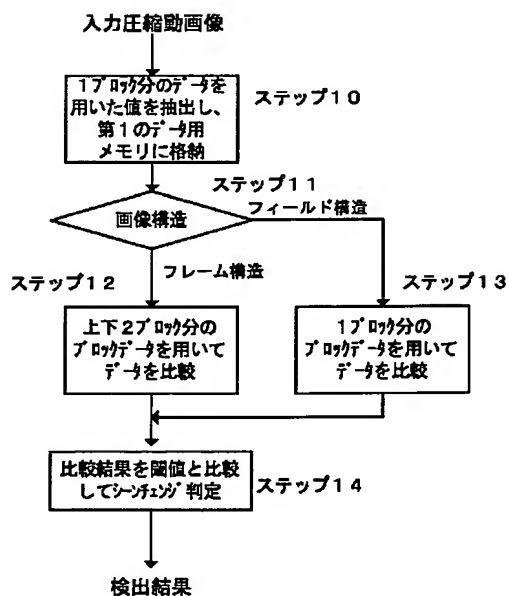
【図6】



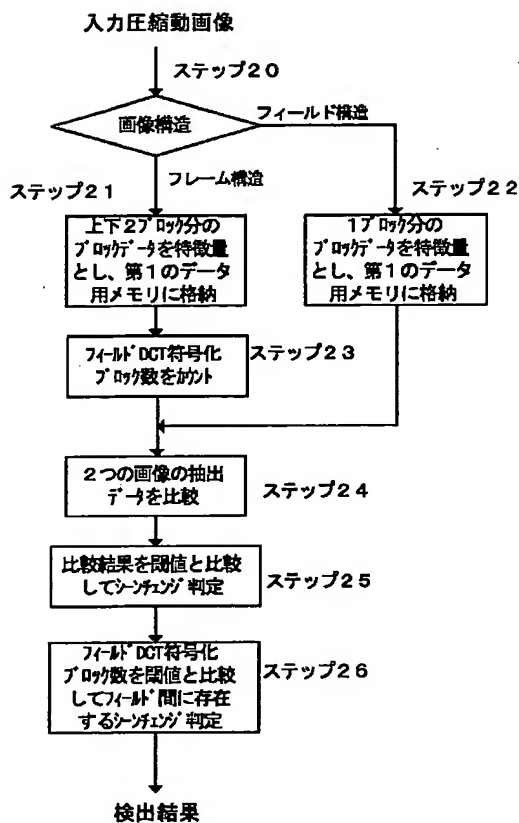
【図7】



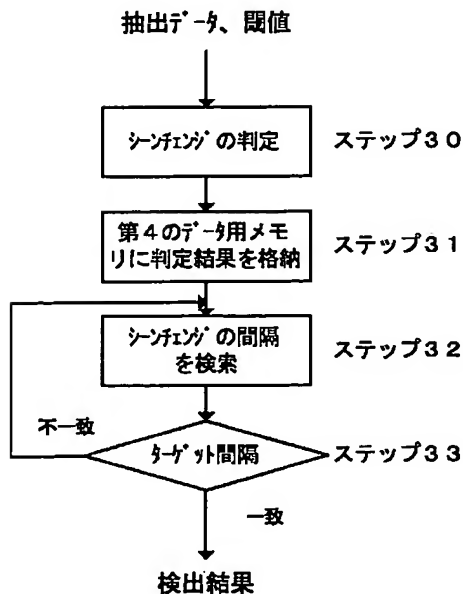
【図8】



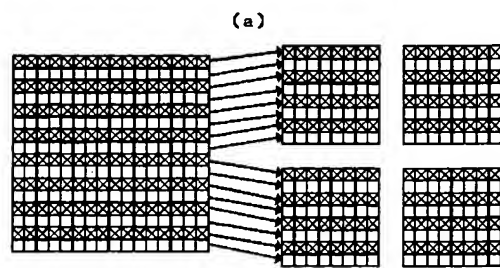
【图 10】



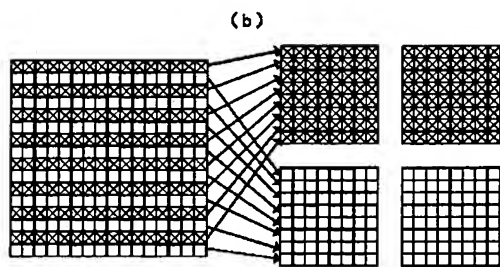
【图 13】



【図15】



フル-DCT符号化におけるマクロブロック構造



Fast-DCT符号化におけるマクロブロック構造

---

フロントページの続き

(72)発明者 池田 淳  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

F ターム(参考) 5C059 KK36 MA00 MA23 NN28 NN43  
PP05 PP06 PP07 TA64 TB07  
TB08 TC14 TC43 TD12 UA38  
UA39  
5J064 AA01 BA16 BB03 BC01 BD02